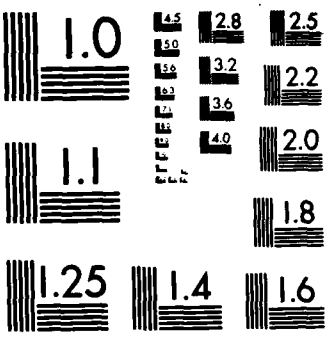
AD-A174 629 FURTHER STUDIES IN ESTIMATION OF LIFE DISTRIBUTION CHARACTERISTICS FROM C (U) SOUTH CAROLINA UNIV COLUMBIA DEPT OF STATISTICS K J PADGETT 02 JUN 86 1/1

UNCLASSIFIED AFOSR-TR-86-2050 AFOSR-84-0156 F/G 12/1 NL



AFOSR-TR- 86-2050

FURTHER STUDIES IN ESTIMATION OF LIFE DISTRIBUTION CHARACTERISTICS FROM CENSORED DATA: ANNUAL TECHNICAL REPORT (6-1-85 to 5-31-86)

W. J. Padgett, Principal Investigator

Grant Number AFOSR-84-0156

Department of Statistics
University of South Carolina
Columbia, SC 29208

June 2, 1986

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
University of South Carolina		AFOSR/NM

6c. ADDRESS (City, State and ZIP Code)	7b. ADDRESS (City, State and ZIP Code)
Department of Statistics Columbia, South Carolina 29208	Building 410 Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR	NM	AFOSR-84-0156

8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
Building 410 Bolling AFB, DC 20332-6448	6.1102F	2304	A5	

11. TITLE (Include Security Classification) Further Studies in Estimation of Life Distribution Characteristics from Censored Data: Annual Technical Report

12. PERSONAL AUTHOR(S)
W. J. Padgett

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Interim	FROM 6-1-85 TO 5-31-86	1986 June 2	13

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Nonparametric estimation; density estimation; Kernel methods; Quantile estimation; Percentiles; Bandwidth; Right-censored data; Asymptotic behavior.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The main objectives of this research have been the development of smooth nonparametric estimators of quantile functions from right-censored data and the further study of smooth density estimators from censored observations. In particular, kernel-type quantile estimators have been obtained under censoring which give better estimates of percentiles of the lifetime distribution than the usual product-limit quantile estimator. During the past year, asymptotic properties of these kernel quantile estimators have been developed, including asymptotic normality, consistency, and mean square convergence. In addition, a data-based procedure for selecting the bandwidth has been investigated using the bootstrap, and approximate confidence intervals for the true quantile have been proposed using bootstrap estimates of the sampling distribution. Theoretical results on the optimal bandwidth selection for kernel density estimators under random right censorship have also been obtained.

New results in several other problem areas were also developed. These included the study of linear empirical Bayes estimators, prediction intervals for the inverse Gaussian

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Maj. Brian W. Woodruff	(202) 767-5027	AFOSR/NM

distribution, nonparametric hazard rate estimation under censoring, nonparametric inference for step-stress accelerated life tests under censoring, discrete failure models, simultaneous confidence intervals for pairwise differences of normal means, and optimal designs for comparing treatments with a control.

1. Summary

This document reports the work performed and other research activities of the investigators during the funding period from June 1, 1984, to May 31, 1986, under grant number AFOSR-84-0156. The results of the past year, June 1, 1985, to May 31, 1986, are emphasized.

The main objectives of this research have been the development of smooth nonparametric estimators of quantile functions from right-censored data and the further study of smooth density estimators from censored observations. In particular, kernel-type quantile estimators have been obtained under censoring which give better estimates of percentiles of the lifetime distribution than the usual product-limit quantile estimator. During the past year, asymptotic properties of these kernel quantile estimators have been developed, including asymptotic normality, consistency, and mean square convergence. In addition, a data-based procedure for selecting the bandwidth has been investigated using the bootstrap, and approximate confidence intervals for the true quantile have been proposed using bootstrap estimates of the sampling distribution. Theoretical results on the optimal bandwidth selection for kernel density estimators under random right censorship have also been obtained.

New results in several other problem areas were also developed. These included the study of linear empirical Bayes estimators, prediction intervals for the inverse Gaussian distribution, nonparametric hazard rate estimation under censoring, nonparametric inference for step-stress accelerated life tests under censoring, discrete failure models, simultaneous confidence intervals for pairwise differences of normal means, and optimal designs for comparing treatments with a control. Work is continuing on smooth nonparametric function

estimation, including the study of generalized quantile function estimates under censoring and the development of a theory of optimal bandwidth selection for the kernel-type quantile estimator based on mean squared error or other criteria. Exact asymptotic and small-sample confidence intervals for quantiles are also under study. All of these results should prove fruitful in the assessment of reliability and maintenance policies of various types of equipment.

2. Research Objectives During the Reporting Period.

The research objectives of this project can be divided into the following categories:

A. Nonparametric estimation of quantiles under censoring;

B. Kernel density estimation from randomly right-censored data;

C. Nonparametric estimation from step-stress accelerated life tests; and

D. Linear empirical Bayes estimation of mean time to failure.

The specific research problems that were considered during this reporting period in these areas will be briefly outlined below. The results obtained to date will be described in Section 3 of this report

A. Nonparametric estimation of quantiles under censoring. One of the main objectives of this research was to further develop properties of smooth nonparametric estimates of the quantile function from randomly right-censored data. The kernel-type quantile estimator proposed by Padgett [paper no. 5 in Section 4] was to be studied with respect to its small-sample, as well as asymptotic, properties. This estimator smooths the product-limit quantile function $\hat{Q}_n(p) = \inf\{t: \hat{F}_n(t) \geq p\}$, where $\hat{F}_n$ denotes the product-limit estimator of the lifetime distribution F_o, and is better than $\hat{Q}_n(p)$

in the sense of smaller mean squared errors in many situations. Thus, such estimators can provide more accurate estimates of percentiles, aiding in the assessment of equipment reliability.

B. Kernel density estimation from randomly right-censored data. Another objective of this project during the reporting period was to determine a method of "optimal" bandwidth selection for kernel-type density estimators under random right-censorship, with respect to some optimality criterion. The kernel density estimators of Blum and Susarla (Multivariate Analysis-V, 1980) and Földes, Rejtö and Winter (Periodica Mathematica Hungarica, 1980) were to be studied.

C. Nonparametric estimation from step-stress accelerated life tests. Usually in accelerated life tests, items are put on test at several accelerated stresses and observed until failure or until the test terminates. A more realistic method is to place items on test and allow the stress levels to change at preassigned times. This will result in more failures in general. The objective of the research in this area was to obtain a strongly consistent nonparametric estimator of the unknown life distribution at the normal use stress. A rescaling technique was utilized to study this problem.

D. Linear empirical Bayes estimation of mean time to failure. In estimating the mean lifetimes of n independent problems, it is hoped that the linear empirical Bayes estimators will converge to the optimal linear Bayes estimators with reasonable rapidity. This convergence was to be studied here. Also, another objective in this area was to identify a natural class of estimators between the linear Bayes estimators and the Bayes estimators.

3. Status of the Research Effort

In this section a substantive statement of the progress and significant accomplishments towards achieving the research objectives for this reportin period outlined in section 2 will be given. The specific research papers referred to will be listed in section 4 of this report, which shows the cumulative results of the entire research project.

A. Nonparametric estimation of quantiles under censoring. Based on right-censored data from a lifetime distribution F_o, a kernel-type estimator of the quantile function $Q^o(p) = \inf\{t: F_o(t) \geq p\}$, $0 \leq p \leq 1$, was proposed by Padgett in paper number [5] listed in section 4. The estimator is defined by $Q_n(p) = h_n^{-1}\int_0^1 \hat{Q}_n(t)K((t-p)/h_n)dt$, which is smoother than the usual product-limit quantile function $\hat{Q}_n(p) = \inf\{t: \hat{F}_n(t) \geq p\}$, where $\hat{F}_n$ denotes the product-limit estimator of F_o. Under the random censorship model and general (but nonrestrictive) conditions on h_n, K, and F_o, it was shown in [5] that $Q_n(p)$ is strongly consistent. In addition, it was shown that an approximation, Q_n^*, to Q_n is asymptotically equivalent to it with probability one. The general conditions for consistency included that $h_n \to 0$ as $n \to \infty$, that K be a symmetric probability density with finite support, and that F_o satisfy certain continuity conditions.

During the past year, the estimator $Q_n(p)$ was studied further and in much more detail. In paper [7] it was shown that, under conditions similar to those required for strong consistency, $n^{\frac{1}{2}}[Q_n(p) - Q^o(p)]$ is asymptotically normally distributed, provided $n^{\frac{1}{4}}h_n \to 0$. This condition on h_n can be replaced by $n^{\frac{1}{2}}h_n \to 0$ by using a slightly different proof. By centering with the quantity $Q^o(p,h) = h_n^{-1}\int_0^1 Q^o(t)K((t-p)/h_n)dt$, an asymptotic normality result can be obtained without the above condition on the rate of convergence of h_n to zero.

Also, using this type of centering, the asymptotic normality of the approximation Q_n^* was shown. These last two results are contained in paper number [12]. In addition, in [12], mean convergence of $Q_n(p)$ and $Q_n^*(p)$ were obtained and the mean square convergence was shown. In particular, an upper bound on the rate of mean square convergence of $Q_n(p)$ to $Q^o(p)$ was shown to be $g(n,h_n) = O(h_n + h_n^{-1/2}n^{-5/4} + (\log n/n)^{3/4} + n^{-1/4}c(1-p)^2)$, where c is a constant. An example of a bandwidth sequence satisfying $g(n,h_n) \to 0$ as $n \to \infty$ is $h_n \simeq c_n n^{-\delta}$ for $0 < \delta < 5/2$ where $c_n > 0$ is a bounded sequence. The value of h minimizing $g(n,h)$ was found to be of the form $h_n = Cn^{-5/6}$. However, this rate is not sharp and further investigation with respect to the asymptotically optimal choice of h (optimal in the sense of minimum mean squared error) is needed.

Since the exact mean squared error has not been calculated, an expression for "optimal" h_n minimizing the mean squared error is not available. Hence, a study of the effect of varying the bandwidth was done by Monte Carlo simulations in paper [15]. Five common life distributions, two different censoring distributions, and three different kernel functions were used in the extensive simulations. The results indicate that at fixed p the kernel estimator $Q_n(p)$ has smaller mean squared error than the product-limit quantile estimator $\hat{Q}_n(p)$ for a range of values of the bandwidth h_n. In addition, in [15] a method of selecting an "optimal" bandwidth value, in the sense of small estimated mean squared error, based on the bootstrap, was investigated. The results were consistent with the simulation study. The bootstrap was also used to obtain interval estimates of $Q^o(p)$ after determining a value of h_n to use in computing $Q_n(p)$.

B. Kernel density estimation from randomly right-censored data.

The problem of asymptotically optimal selection of the bandwidth sequence for kernel density estimators under censoring was studied in paper number [13]. The usual kernel density estimator $f_n(x) = h_n^{-1}\int_{-\infty}^{\infty} K((x-t)/h_n)d\hat{F}_n(t)$ and the Blum-Susarla estimator given by $f_n^*(x) = [nh_nH_n^*(x)]^{-1}\ \Sigma_{j=1}^{n} K((x-X_j)/h_n)\cdot I_{[\Delta_j=1]}$ were considered, where H_n^* denotes an estimator of the censoring survival function, (X_j,Δ_j), $j=1,\ldots,n$, denote the right-censored data, and $I_{[\cdot]}$ denotes the indicator function of the event $[\cdot]$. Two important contributions to the theory of bandwidth selection for kernel density estimators under right censorship were made in this paper. First, an asymptotic representation of the integrated squared error of each estimator in terms of easily understood variance and squared bias components is given. Second, it is shown that if the bandwidth is chosen by the data-based method of least squares cross-validation, then it is asymptotically optimal in a compelling sense. A by-product of the first part is an interesting comparison of the two kernel density estimators which shows that the difference between the two is typically not negligible.

C. Nonparametric estimation from step-stress accelerated life tests.

In accelerated life testing, items are subjected to greater stress than that under the normal use conditions and, from the resulting failure data, an estimate of the lifetime distribution (or other inferences) under the normal use conditions is obtained. Most of the published work on inference from accelerated life tests require at least one of the following restrictions: (i) that every item which is subjected to an overstress is observed until it fails or is removed (censored) under constant application of the stress, or (ii) that the lifetime distribution of the items under every accelerated

stress level is assumed to be known except for the underlying parameters. A more realistic approach allows the stress on an unfailed item to change at a preassigned test time. This approach is used in government defense studies and is referred to as step-stress testing.

In paper number [11] step-stress testing is considered from a nonparametric point of view when the observations have been arbitrarily right censored, relaxing both restrictions (i) and (ii) above. Thus, no assumptions are made about the particular form of the underlying lifetime distribution. Using a rescaling method, an estimator of the life distribution under the normal use conditions is obtained. The estimator is shown to be strongly consistent.

D. Linear empirical Bayes estimation of mean time to failure. In paper [10], it has been successfully established that not only the linear empirical Bayes estimator approaches the optimal linear estimator with probability one, but also in the quadratic mean with a rate of convergence inversely proportional to the number of problems considered. The same is true for the mean squared error of the estimators. Indeed, the limiting constant has been explicitly evaluated, which will provide a guideline for the consideration of a moderate number of problems. Effort will be continued to study the moderate sample properties of the linear empirical Bayes estimators, and to identify a natural and appropriate intermediate class of estimators.

E. Miscellaneous. In other work, an invited paper [4] was written on nonparametric methods for hazard rate estimation from censored data. This paper first reviewed the available methods in the literature and then presented some new results on a kernel-type estimator of the hazard rate under random right censoring. The Koziol-Green (or proportional hazards) model was assumed for asymptotic results on the bias and mean squared error of the kernel estimator.

In paper number [9], a comparison was given of two prediction intervals for a future observation from the inverse Gaussian distribution. This distribution arises as a lifetime model which is an alternative to the lognormal distribution. One prediction interval is exact and the other is an approximate one. Surprisingly, the comparison showed that the approximate interval is superior in general, since it always yields a two-sided interval and has high coverage probability as well as small expected length.

In paper number [14] a regenerative sampling plan is developed for the sequential comparison of two populations having positive integral response. It is designed to be both an extension and an improvement of the play-the-winner rules for binary trials in the sense that a much wider variety of responses is allowed, the fraction of inferior selections approaches zero, and the play-the-winner rule is contained as a special case. Almost sure convergence and moment convergence in the pth order is studied for the fraction of inferior selections and for a maximum likelihood estimator of the mean response. A conditional test of hypothesis is given for the binary case. These results have applications in clinical trials and in quality control as well as in a variety of other sequential situations.

4. Cumulative List of Written Publications in Technical Journals

In this section, the research papers that have been written under this grant are listed. They are divided into four categories: In print, accepted for publication, and in progress. Copies of the manuscripts and reprints have been forwarded to the Program Manager as they were submitted or published.

A. In Print

[1] André M. Lubecke and W. J. Padgett, Nonparametric maximum penalized likelihood estimation of a density from arbitrarily right-censored observations. Communications in Statistics-Theory & Methods 14 (1985), 257-271.

[2] J. D. Spurrier and S. P. Isham, Exact simultaneous confidence intervals for pairwise comparisons of three normal means. Journal of the American Statistical Association 80 (1985), 438-442.

[3] W. J. Padgett and J. D. Spurrier, On discrete failure models. IEEE Transactions on Reliability R-34 (1985), 253-256.

[4] D. T. McNichols and W. J. Padgett, Nonparametric methods for hazard rate estimation from right-censored samples. Journal of the Chinese Statistical Association 23 (1985), 1-15.

[5] W. J. Padgett, A kernel-type estimator of a quantile function from right-censored data. Journal of the American Statistical Association 81 (1986), 215-222.

[6] J. D. Spurrier and D. Edwards, An asymptotically optimal subclass of balanced treatment incomplete block designs for comparisons with a control. Biometrika 73 (1986), 191-199.

B. Accepted for Publication

[7] Y. L. Lio, W. J. Padgett and K. F. Yu. On the asymptotic properties of a kernel-type quantile estimator from censored samples. Journal of Statistical Planning and Inference (to appear).

[8] W. J. Padgett, Nonparametric estimation of density and hazard rate functions when samples are censored. Handbook of Statistics, Vol.7, P. R. Krishnaiah, Ed. (to appear).

[9] W. J. Padgett and S. H. Tsoi, On prediction intervals for future observations from the inverse Gaussian distribution. IEEE Transactions on Reliability (to appear).

[10] K. F. Yu, On the bounded regret of empirical Bayes estimators. Communications in Statistics-Theory and Methods (to appear).

C. Submitted for Publication

[11] D. T. McNichols and W. J. Padgett, Inference for step-stress accelerated life tests under arbitrary right-censorship. Submitted to Journal of Statistical Planning and Inference.

[12] Y. L. Lio and W. J. Padgett, Some convergence results for kernel-type quantile estimators under censoring. Submitted to Statistics and Probability Letters.

[13] J. S. Marron and W. J. Padgett, Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples. Submitted to Annals of Statistics.

[14] S. D. Durham and K. F. Yu, Regenerative sampling and monotonic branching processes. Submitted to Annals of Statistics.

[15] W. J. Padgett and L. A. Thombs, Smooth nonparametric quantile estimation under censoring: Simulations and bootstrap methods. Submitted to Communications in Statistics-Simulation and Computation.

D. Research in Progress

[16] Y. L. Lio and W. J. Padgett, Generalized product-limit quantile estimators.

[17] K. F. Yu, A linear regression with unobserved dependent variables.

5. Professional Personnel Associated with the Research Effort

In addition to the principal investigator, W. J. Padgett, several co-investigators have been partially supported by this grant since June 1, 1984. John D. Spurrier was partially supported during the period from June 1, 1984, to May 31, 1985, and Kai F. Yu was partially supported from June 1, 1985, to May 31, 1986. During the summer of 1985, Diane T. McNichols, Department of Statistics, Virginia Polytechnic Institute and State U., was partially supported. Also, two graduate students have been supported, André M. Lubecke during the summer of 1984, and Y. L. Lio from August 16, 1984, to May 31,1986. Ms. Lubecke completed the Ph.D. degree in August, 1985, and Mr. Lio is currently working on his dissertation, having completed all written doctoral exams.

6. Interactions

The investigators attended several meetings and conferences and gave (invited and contributed) talks as follows:

June 1, 1984, to May 31, 1985:

i) W. J. Padgett, J. D. Spurrier attended the Conference on Reliability and Quality Control, University of Missouri, Columbia, MO, June 4-8, 1984.

ii) W. J. Padgett (joint with A. M. Lubecke), "Nonparametric maximum penalized likelihood estimation of a density from arbitrarily right-censored data." Annual Meeting of IMS in Tahoe City, CA, August 21-24, 1984.

iii) J. D. Spurrier, "The training of statistical consultants at the University of South Carolina" (invited). Joint Statistical Meetings, Philadelphia, PA, August, 1984.

iv) D. Edwards, J. D. Spurrier, "An asymptotically optimal subclass of BTIB designs for comparisons with a control." Joint Statistical Meetings, Philadelphia, PA, August, 1984.

v) J. D. Spurrier and W. J. Padgett, "On discrete failure models." Spring Regional Meeting of the Biometric Society (ENAR) and ASA in Raleigh, NC, March 25-27, 1985.

vi) W. J. Padgett was the invited discussant for the Session on Reliability and Life Testing, Spring Regional Meeting of the Biometric Society (ENAR) and ASA in Raleigh, NC, March 25-27. 1985.

vii) W. J. Padgett, "Some estimation procedures for quantile and density functions from censored data." AFOSR Workshop on Reliability, Luray, VA, May 29-31, 1985.

June 1, 1985, to May 31, 1986:

i) W. J. Padgett gave an invited talk on "Nonparametric density and failure rate estimators when samples are censored" at the SREB-ASA Summer Research Conference on Statistics, Boone, NC, June 16-21, 1985.

ii) W. J. Padgett, "A kernel type estimator of a quantile function from right-censored data." Joint Statistical Meetings, Las Vegas, NV, August 5-8, 1985.

iii) K. F. Yu. "A note on the bounded regret of empirical Bayes estimators," 191st Meeting of the IMS, Stoney Brook, NY, June 12-13, 1985.

iv) K. F. Yu, "On the bounded regret of empirical Bayes estimators." Joint Statistical Meetings, Las Vegas, NV, August 5-8, 1985.

v) K. F. Yu, "On a renewal theorem." Centenary Session of the ISI, Amsterdam, The Netherlands, August, 1985.

vi) K. F. Yu organized and chaired a session on Adaptive Statistics at the Symposium on Adaptive Statistical Inference and Related Topics, Brookhaven National Laboratory, NY, June 8-11, 1985.

vii) K. F. Yu attended the NSF-CBMS Conference on Nonparametric Priors, University Park, PA, June 3-7, 1985.

viii) W. J. Padgett (with Y. L. Lio), "Some convergence results for kernel-type quantile estimators from censored data," Spring Statistics Meetings (ASA/ENAR), Atlanta, GA, March 16-19, 1986.

ix) W. J. Padgett gave an invited talk on "Some results for kernel-type quantile estimators from right-censored data" at the U.S. Army Research Office Workshop on Life Testing, Washington, DC, March 9-12, 1986.

x) W. J. Padgett and K. F. Yu attended the 8th Annual Meeting of the S.C. Chapter of ASA, Columbia, SC, April 25, 1986.

xi) K. F. Yu attended the Spring Statistics Meetings, Atlanta, GA, March 16-19, 1986.

7. Inventions, Patent Disclosures, and Applications Stemming from the Research Project

No inventions or patents have stemmed from this research.

The results reported in Section 4 have wide application in the estimation and assessment of reliability and maintenance of military equipment. The various nonparametric procedures with censored data developed in this project allow accurate estimation of median and mean lifetime, percentiles, densities, survival probability, and other lifetime characteristics without assuming particular forms of the life distribution. In particular, the kernel-type quantile estimator will provide more accurate estimates of the median lifetime and smaller percentiles of a piece of equipment than the usual product-limit estimators.

8. Other Professional Activities

During this reporting period, June 1, 1985, to May 31, 1986, the investigators have been involved in numerous other professional activities that are intimately related to the research efforts on this grant. Professor Padgett has refereed eleven manuscripts, 2 for Technometrics, 3 for the Journal of the American Statistical Association, 2 for Communications in Statistics, 1 for Statistics and Probability Letters, 1 for Pakistan Journal of Statistics, and

1 for the Journal of the University of Kuwait (Science). Professor Yu refereed one article for the Annals of Statistics and one for Sequential Analysis. In addition, 5 papers were reviewed for the Zentrablatt für Mathematik and 14 were reviewed for the Mathematical Reviews. Also, one research proposal was reviewed for the AFOSR and one for the ARO.

Professor Padgett was a member of the International Editorial Board of the Communications in Statistics-Theory and Methods and was an Associate Editor of the Journal of Statistical Computation and Simulation. In addition, he was the Program Co-Chair for the SRCOS-ASA Summer Research Conference in Statistics to be held in Mobile, AL, on June 15-20, 1986. He also chaired an invited paper at the Spring Statistics Meetings in Atlanta, GA, in March, 1986.

END

1-87

DTIC